

Conference ESCAPE to the Future | 25-26 October 2022

Royal Belgian Institute of Natural Sciences | Brussels, Belgium

### 25 October 2022, 11:35 - 12:30 ESCAPE OSSR - Foster collaboration to create FAIR software for science



Open-source Scientific Software and Service Repository

Kay Graf ECAP, Friedrich-Alexander University Des Small JIVE



Mohammad Al-Turany FAIR, GSI Helmholtz



Matthias Fueßling CTAO



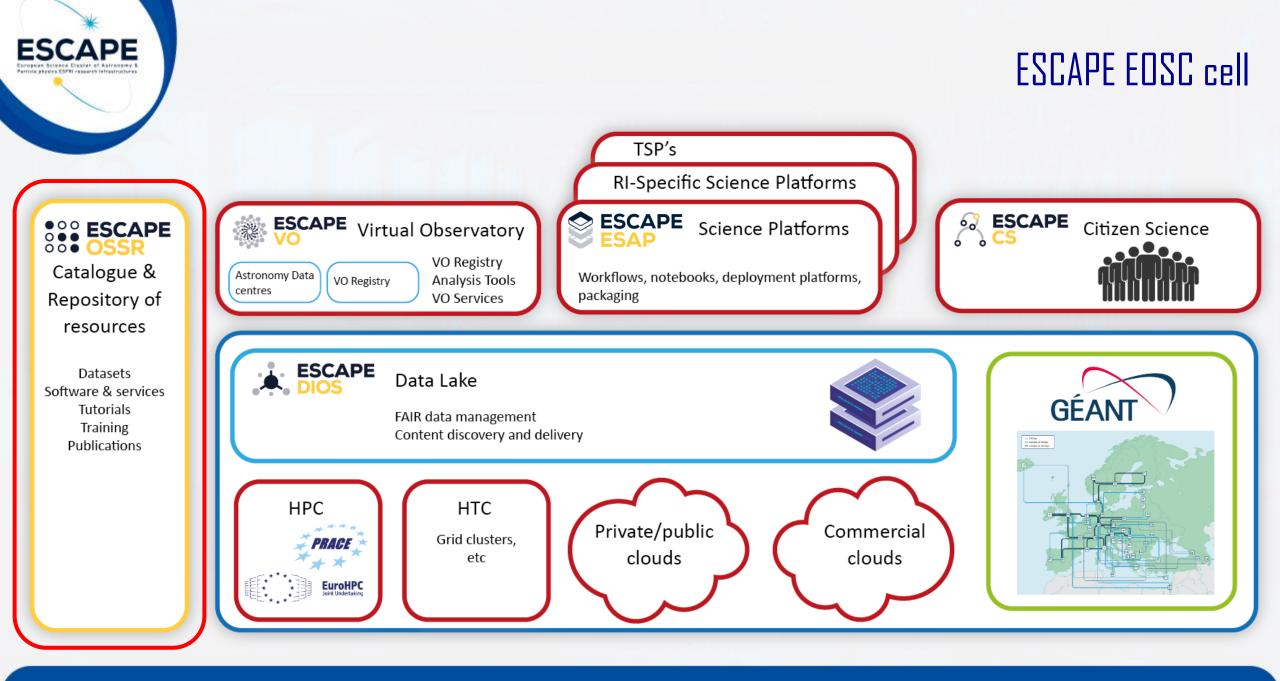
ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures **ESCAPE to the Future** 25-26 October 2022 Brussels, Belgium

# ESCAPE to the Future: OSSR – Open Software in the EOSC

### Kay GRAF

for the OSSR Team ESCAPE to the Future, Brussels, 25-26<sup>th</sup> of Oct. 2022





E-OSSR, ESCAPE to the Future, K. Graf



...

## Status before ESCAPE and OSSR Vision

Software second to data in the EOSC scheme

Oiverse status of (open) scientific software within the community

Cross-experiment analysis via MoUs with minimal exchange of data and software

- Modern approaches individually handled, no extensive cross-fertilisation
- No standard for metadata and archival, no link (of community software) to EOSC

The ESCAPE Open-source Scientific Software and Service Repository (OSSR) is a reliable, sustainable openaccess repository to share scientific software and services to the science community and enable open science.

It houses astro-particle-physics-related scientific software and services for data processing and analysis, as well as test data sets, user-support documentation, tutorials, presentations and training activities.

## **<u>OSSR</u>** Aims and Objectives

#### **Objectives:**

**ESCAPE** 

Catalogue &

Repository of

resources

Datasets Software & services

> Tutorials Training

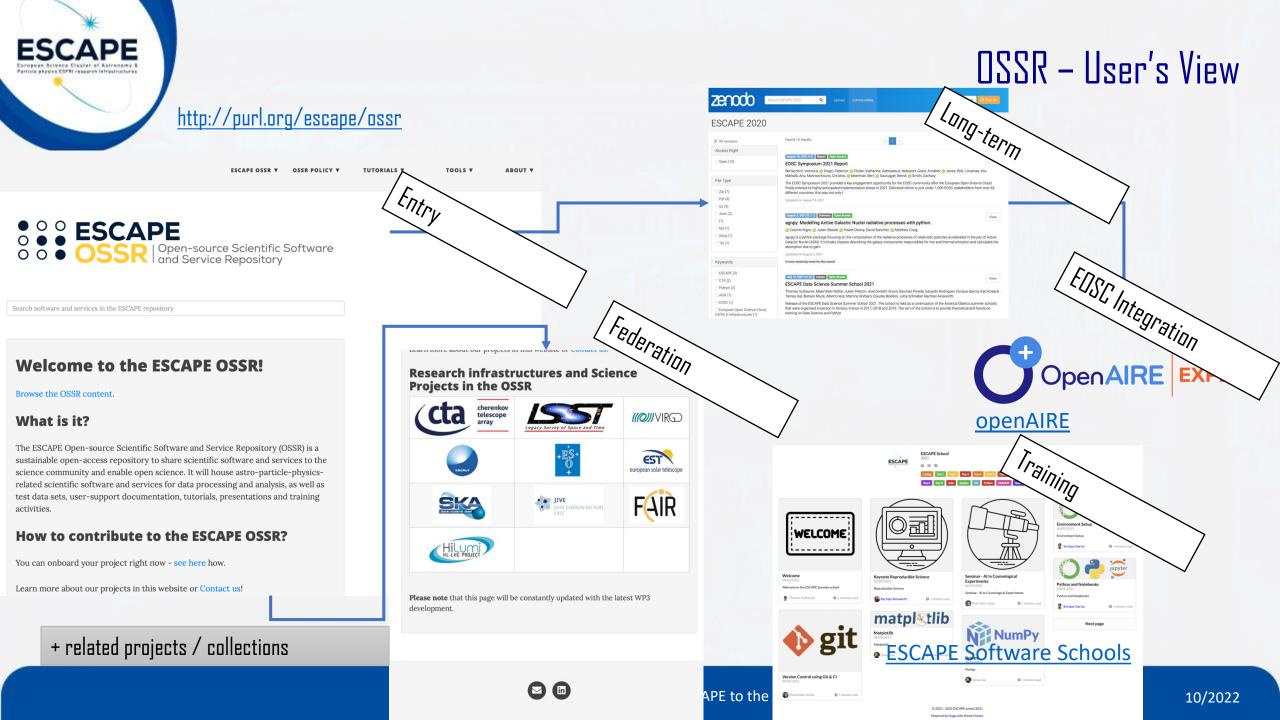
Publications

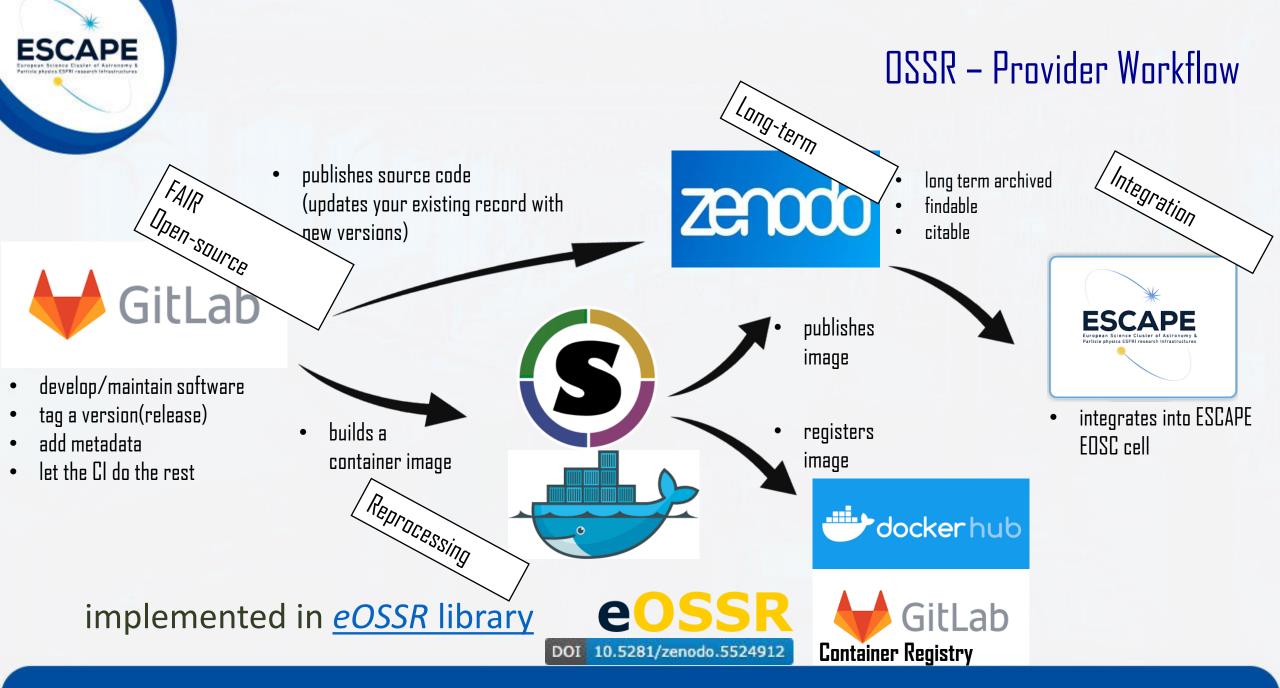
000

- Facilitate and support continuous development, deployment, exposure and preservation of partners' software/tools/services
- Foster interoperability, software re-use and cross-fertilisation between ESFRIs (e.g. simulation)
- Offer an open innovation environment for open standards (e.g. workflows, data-formats), common regulations and shared (novel) software for multi-messenger & multi-probe data
- Establish the link of the community to the EOSC and vice-versa.
- Training of experiences code custodians to create and maintain high-quality, open software and raise their visibility

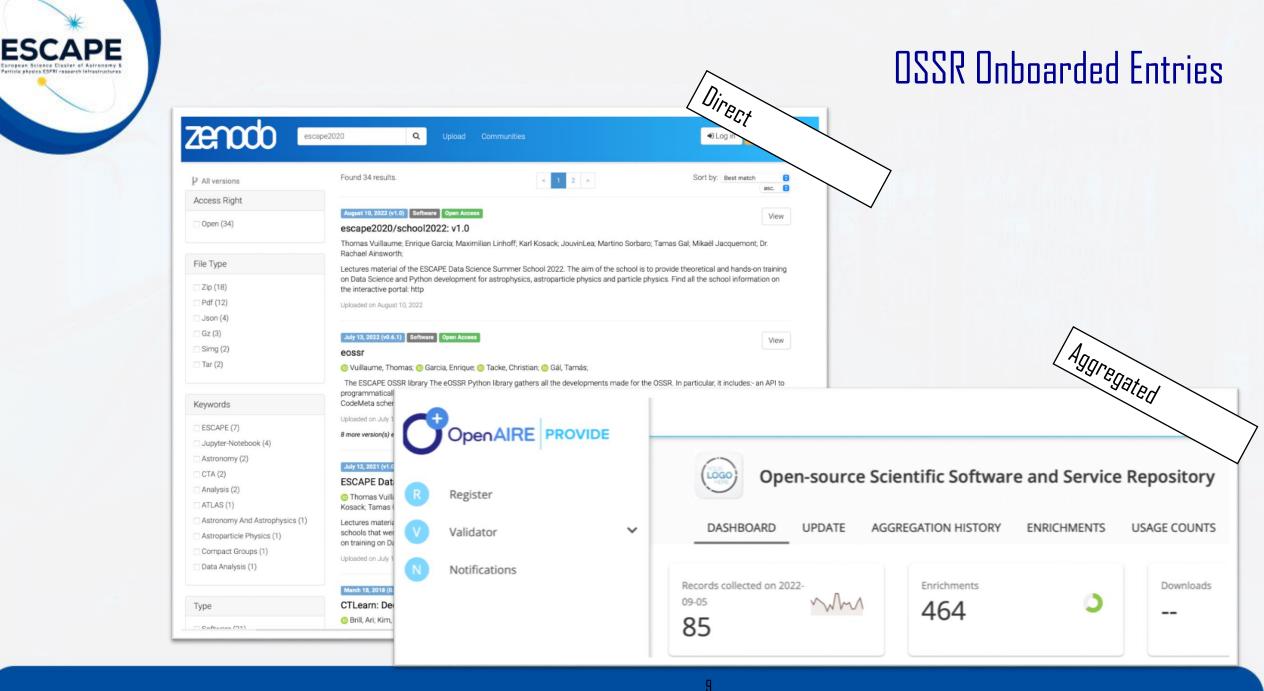
#### All objectives follow:

- Paradigm of enabling open science with software as "first class citizen"
- a **community-based and inclusive** approach
- the FAIR principles for open science resources software and derivatives
- Federation of available resources





#### E-OSSR, ESCAPE to the Future, K. Graf





## Cross-fertilisation and ESFRI Software Developments



**Des Small** JIV-ERIC *CASA improvements for VLBI* 

> **Mohammad Al-Turany** FAIR *DSSR at GSI/FAIR*





**Matthias Füßling** CTA *Cheronkov Telescope Array Observatory (CTAO)* 

E-OSSR, ESCAPE to the Future, K. Graf

#### ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

## **ESCAPE to the Future** 25-26 October 2022 Brussels, Belgium

# CASA improvements for VLBI Des Small, JIVE





## JIVE and the EVN



Image by Paul Boven (boven@jive.eu). Satellite image: Blue Marble Next Generation, courtesy of Nasa Visible Earth (visibleearth.nasa.gov).





### JIVE and the EVN



- JIVE is a European Research InfrastructureWorks for EVN
- Correlates most EVN experiments
- Maintains software correlator (SFXC)
- Supports users in data reduction
  - Originally AIPS
  - Increasingly CASA

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

ESCAPE





Title Here







ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

26/10/2022

### CASA

APE



- Data reduction for Radio Astronomy
- Developed by NRAO, ESO, NAOJ & JIVE
  - C++ and Python
- Handles full data reduction
- Built for connected element arrays
- Used by VLA and ALMA pipelines
- Development process well defined



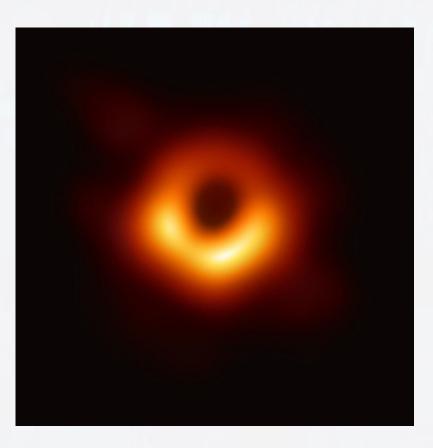
## Each feature is a JIRA ticket and git branch

- Scope of the issue defined
- Implementation, verification, validation
  - Verification requires unit tests and regression tests
  - Validation done by someone else
  - Centralised package building/testing on commit (can be turned on)
- Software engineering best practices!
- (We really don't want to break ALMA pipelines!)



### CASA, VLBI and JIVE

CASA didn't have VLBI features
NRAO always wanted them
JIVE added them
Continued work on new features





ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

ESCAPE



## New CASA papers for added FAIRness!

the S

Search.

Help | Advanc

# Cornell University

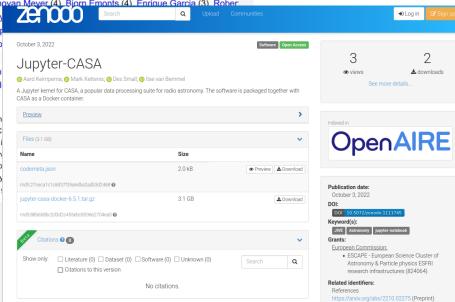
Astrophysics > Instrumentation and Methods for Astrophysics

#### [Submitted on 5 Oct 2022]

#### CASA, the Common Astronomy Software Applications for Radio Astronomy

THE CASA TEAM, Ben Bean (1), Sanjay Bhatnagar (2), Sandra Castro (3), Jennifer Donvan Meyer (4) Bjorn Emonts (4). Enrique Garcia (3), Rober Garwood (4), Kumar Golap (2), Justo Gonzalez Villalba (3), Pamela Harris (2), Yohei Hay Jagannathan (2), Wataru Kawasaki (5), Aard Keimpema (6), Mark Kettenis (6), Jorge Lop McNichols (4), David Mehringer (4), Renaud Miel (5), George Moellenbrock (2), Federico Petry (3), Martin Pokorny (2), Ryan Raba (4), Urvashi Rau (2), Darrell Schiebel (4), Neal (5), Des Small (6), Jan-Willem Steeb (4), Kanako Sugimoto (5), Ville Suoranta (4), Takah (6), Akeem Wells (4), Wei Xiong (1), Arpad Szomoru (6), Morgan Griffith (4), Brian Glend Socorro, (3) ESO, (4) NRAO Charlottesville, (5) NAOJ, (6) JIVE, (7) IDIA)

CASA, the Common Astronomy Software Applications, is the primary data processing software for th Karl G. Jansky Very Large Array (VLA), and is frequently used also for other radio telescopes. The C synthesis, and Very Long Baseline Interferometery (VLBI) telescopes. One of its core functionalities VLA Sky Survey (VLASS), and the Nobeyama 45m telescope. This paper presents a high-level over procedures for calibrating and imaging astronomical radio data in CASA. CASA is being developed b based at the National Radio Astronomical Observatory (NRAO), the European Southern Observatory and the Joint Institute for VLBI European Research Infrastructure Consortium (JIV-ERIC), under the ?



Search ....

Help | Advar



Astrophysics > Instrumentation and Methods for Astrophysics

[Submitted on 5 Oct 2022]

#### CASA on the fringe -- Development of VLBI processing capabilities for CASA

ichael Janssen, George A. Moellenbrock, Dirk Petry, Ciriaco Goddi, Justin D. Linford, Kazi L.J. Ryg dina, Neal Schweigart, Marjolein Verkouter, Aard Keimpema, Arpad Szomoru, Huib Jan van

rometry (VLBI) data has been implemented in the CASA package. This includes two new tasks to handle fringe sting tasks have been adjusted to handle VLBI visibility data and calibration meta-data properly. With these and spectral line observations in CASA. This article describes the development and implementation, and uropean VLBI Network or Very Long Baseline Array data in CASA. Though the CASA VLBI functionality has rizon Telescope data processing, in this paper we compare results for the same dataset processed in CASA s and conclude that CASA in some cases performs better, though it cannot match AIPS for single-core for easy development of pipelines or Jupyter notebooks, and thus contributes to raising VLBI data processing to and reusability.



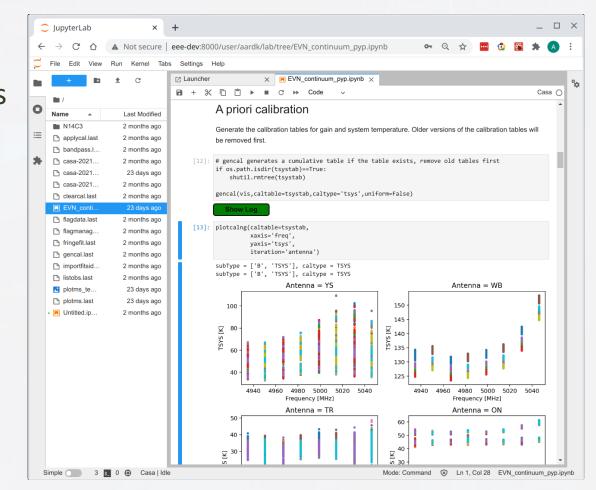
ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

https://arxiv.org/abs/2210.02276 (Preprint)

26/10/2022

## Our OSSR contribution

Wrapped CASA in Jupyter notebooks
We can offer "CASA in the cloud"
Trial run at ERIS 2022
Will be covered by A Keimpema





ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

CAPE



# Thanks!





in company/projectescap@2022

ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures **ESCAPE to the Future** 25-26 October 2022 Brussels, Belgium

# ESCAPE to the Future: OSSR @ GSI/FAIR

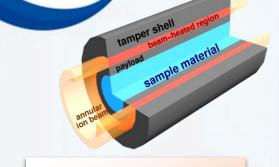
Mohammad Al-Turany GSI Helmholtzzentrum für Schwerionenforschung



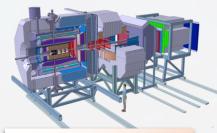








Atomic, applied and plasma physics -APPA

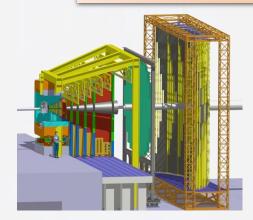


Hadron structure - PANDA

## 1 TByte/s into online farms 35 PByte/year on disk

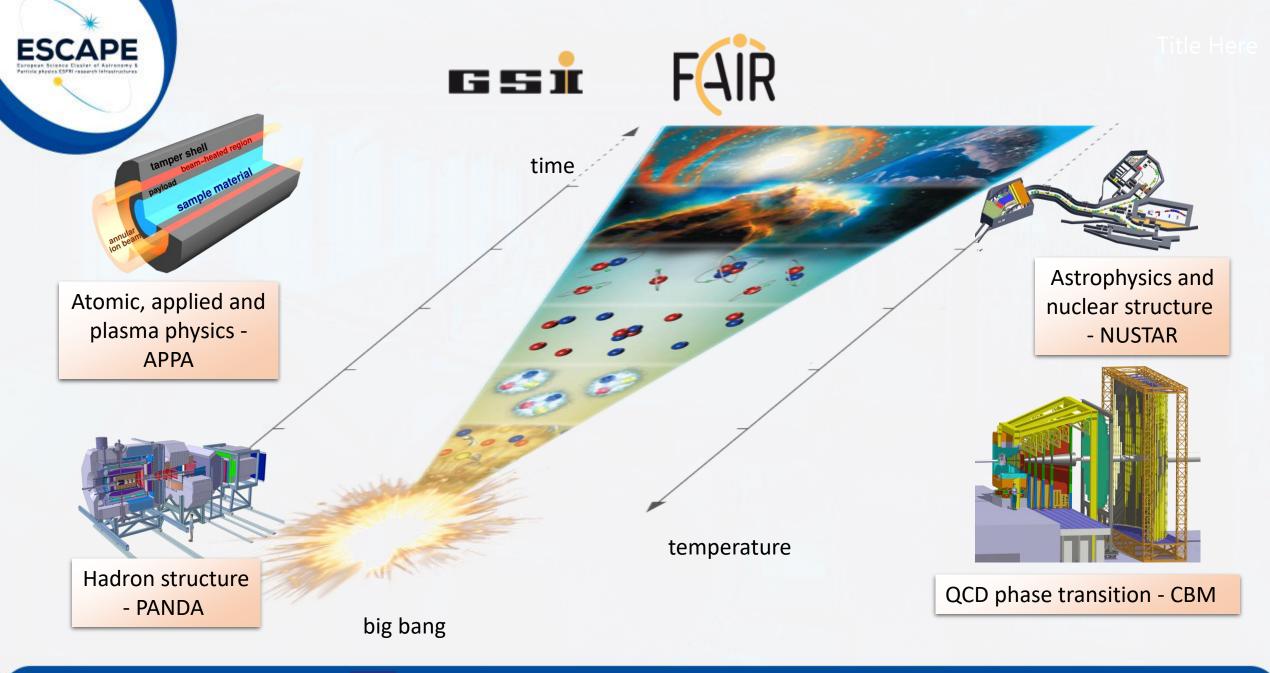


Astrophysics and nuclear structure - NUSTAR



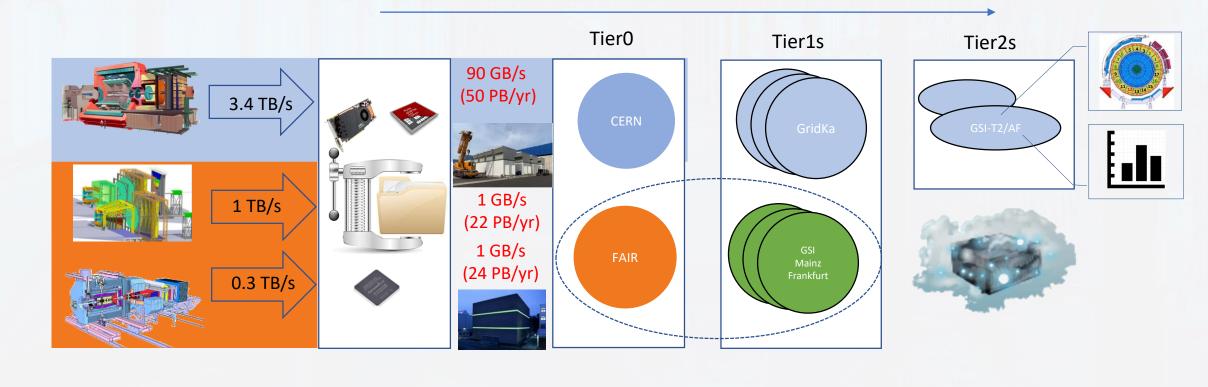
#### QCD phase transition - CBM







## Bring the computing experts from experiments and the IT together to solve common problems





from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

ESCAPE

## ESCAPE

### Research Field Matter

## GSI is member of the Helmholtz association



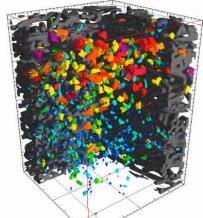
	LK II (User facilities)	International Research Infrastructures	National Research Institutes/ Infrastructures
DESY	FLASH PETRA III IDAF	LHC Belle II CTA (under construction) IceCube European XFEL ESRF	CFEL CSSB NanoLab DESY Test Beams DAF HIB@European XFEL PITZ
FZJ	JCNS (in MLZ)	ESS (under construction) ILL	(FRM-II)
GSI	UNILAC SIS18 ESR	FAIR (under construction) ALICE@LHC	HI Jena HI Mainz
Hereon	GEMS	ESS (under construction)	EMSC
HZB	BESSY II		SupraLab EMIL
HZDR	ELBE HLD IBC	European XFEL EMFL ESRF	HIB@European XFEL DRESDYN
кіт	GridKa	KATRIN Auger IceCube	ATP FLUTE TLK SR Beamlines

#### • Unique Research Facilities

- Many scientific domains and a diverse user community from university, research institutes and industry
- Digitalisation is important for
  - Efficient and sustainable operation
  - Optimum use of research infrastructures
  - Knowledge extraction from research data
  - Frontier science as a driver of innovation
- Topic DMA established in POF IV

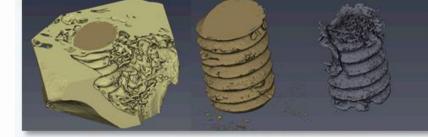






Near real-time **segmentation** of battery electrode data by **AI** 

GSI is leading the Digital Scientific Method topic in Tile Her Matter/Data Management and Analysis (DMA)



Near real-time **segmentation** of bone implant data by **AI** 



**Exascale** simulations of high energy density plasmas

2023: DMA Open Solutions Toolbox

laser-driven ion acceleration

> 2025: Near real-time capabilities

> 2027: Surrogate modelling



Exascale Computing

research in Matter.

High Throughput Computing

Develop, apply and share cutting edge digital

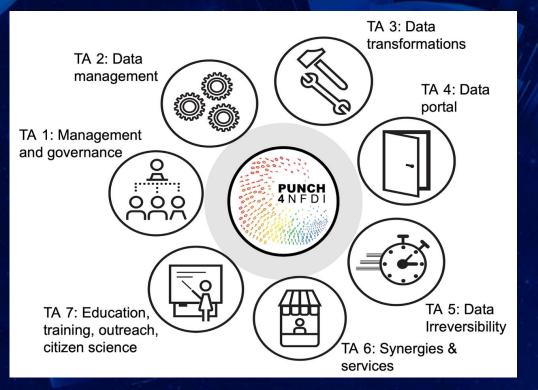
methods and frontier technologies for

- Quantum Computing
- Near real-time analysis

ESCAPE Particle physics ESFRI research Infrastructures

# GSI is member of Punch4NFDI

Deutsche Forschungsgemeinschaft





#### Particles, Universe, NuClei and Hadrons for the NFDI



Nationale Forschungsdaten Infrastruktur

### See also DFG.de/nfdi and nfdi.de



## What do we gain from OSSR?

- Establish modern collection-/link-site with one entry point for software
- Find solutions and environments for workflows rather than services
- Not only the software itself but also the environment that enables the scientific community to use/test the software, e.g. documentation, continuous integration and deployment services and evaluation data sets.





# What have been done up to now?



## What have we done up to now?

### Few Packages are already onboarded:

FairRoot:

A simulation, reconstruction and analysis framework that is based on the ROOT system. The user can create simulated data and/or perform analysis with the same framework.

FairMQ:

C++ Message Queuing Library and Framework

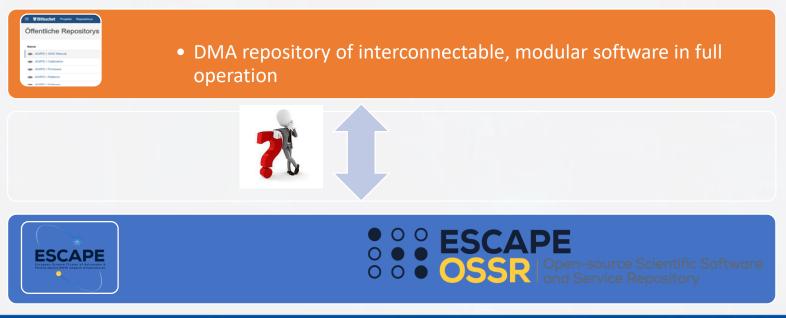
### DDS:

The Dynamic Deployment System (DDS) - is a tool-set that automates and significantly simplifies a deployment of user defined processes and their dependencies on any resource management system using a given topology.



### What have we done up to now?

- The E-OSSR Implements an open, inclusive repository (catalogue) for the Astrophysics, Astroparticle Physics, Particle Physics community
- With DMA we add the Nuclear structural physics and Photon physics in HGF to the club







### What have we done up to now?

Onboarding DMA software to OSSR:

• GSI take over the organization of the onboarding process for DMA software

 First DMA project already onboarded and will serve as an example for the DMA community

R3BRoot:

Software for simulations and data analysis of Reactions with Relativistic Radioactive Beams experiment at FAIR





# What is next?





### What is next:

### Within the DMA community, a list of projects is being prepared for onboarding

### Requirement to OSSR:

 Data should only be part of the repository if supporting to software (e.g. Training data sets for ML algorithms)



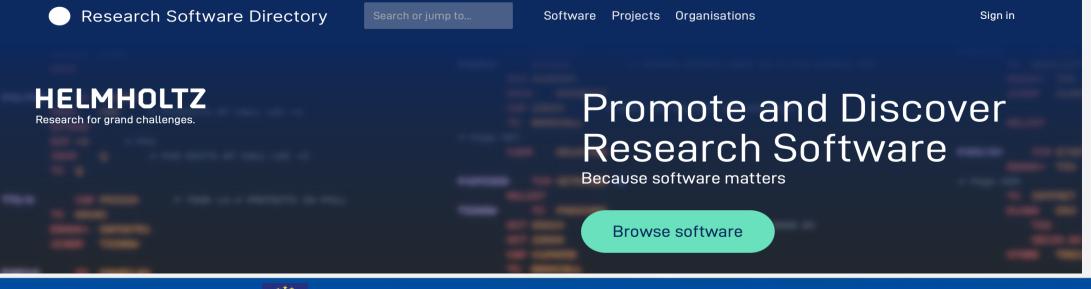


## What is next:



HIFIS Research Software Directory: highly flexible software directory, but no curation

Bring together the Helmholtz Research Software Directory and the OSSR. (https://helmholtz.software/)







# Thanks!





in company/projectescape

ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures **ESCAPE to the Future** 25-26 October 2022 Brussels, Belgium

# Cherenkov Telescope Array Observatory (CTAO)

Matthias Füßling, Gareth Hughes (CTAO)



## CTAO and open research software

- CTAO is an open observatory aimed at the whole science community as CTAO users
- CTAO is committed to support open science and follow FAIR principles
- CTAO acquires raw data, processes and calibrates the raw data to science-ready data as a service task
  - Software for the data processing, calibration and simulation is open
  - Increase chances of sharing of software across ESFRIs and increases credibility of provided data
- CTAO will provide high-quality science-ready data and software to its users
  - Open research software to analyse the data (science analysis tools): from photon lists to sky maps, spectra, light curves
  - High-quality software and user support as important as high-quality data
  - Includes maintenance and feature upgrades of software
  - Includes user support for the analysis of data, with training and schools
- Multi-messenger, Multi-wavelength, Multi-experiment science is one of the cornerstones
  - Aim for interoperability of data products, and science software
  - Exchange of workflows, reproducibility
- Some CTAO software based on open community-based software products
  - Benefit from the community, and give back to the community
  - Increase chances of sharing of software across ESFRIs and communities



## CTAO and OSSR

With upcoming EOSC, users will explore additional ways to approach science

- Usually, the users start with discovery of data
- Users now also discover software and/or analysis workflows

Ousers will have more ways to search and access software and data, and to collaborate

- First point of contact for CTAO-specific data and software: the CTAO science portal
- For cross-ESFRI workflows, OSSR via EOSC may play an important role

### OSSR provides to CTAO

- e a set of standards which CTAO turns into recommendations and guidelines for creating our observatory's software (incl. SW metadata, licensing, provenance, SW on-boarding process)
- e a repository to easily find and access software and workflows
- e a forum for software development and cross-ESFRI fertilization

CTAO benefits from OSSR

Enlarge CTAO user base and use of CTAO data products and software, enable new science



## Software Developments in OSSR

CTAO still under construction

- official software yet
- Community-based software projects for CTA:
  - Aim to extend features, enhance software quality
  - Examples: GammaPy, agnpy, glike, ctlearn, gammalearn, hipeRTA
- Software onboarded to OSSR
  - Findable and useable

### Cross-ESFRI software developments

- Aim for interoperability, multi-instrument, multi-messenger
- Examples: Concordia, Wavefier (see next slides)

### Continued in EOSC-Future

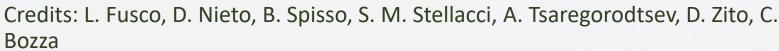
- Aim to test and extend software with Test Science Projects
- Example: MM/MWL analysis pipeline for AGN models (MAP, S. Lloyd)

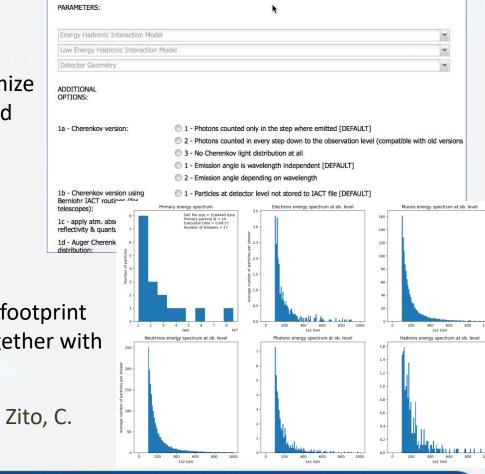
## Concordia – Containers for Corsika on DIRAC

- coordinated work between multiple ESFRIs partners for common software (together with CEVO)
  - Here: KM3NeT and CTAO
- Development and production of CORSIKA turnkey containers for various use cases and functional development of CORSIKA for specific purposes
  - CORSIKA is an air shower simulation program
- Usage of the DIRAC middle-ware for large scale simulation productions is pursued
  - see ESAP session for batch processing

GUI to customize containers and manage jobs

Performance footprint published together with container





41



ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

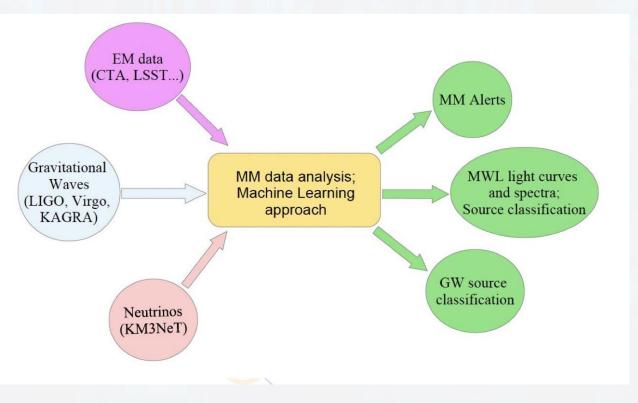
26/10/2022



## Wavefier – an analysis platform for multi-messenger data

#### (see also dedicated talk by Alberto less)

- OSSR contains competence center for software and service innovation (lead by E. Cuoco)
  - Exchange of experience and harmonisation of approaches for innovative workflows between different partners
- Wavefier is an analysis platform for incoming MM data, coordinate follow-up observations, use novel ML-based pipelines for MM analysis
- Here: extend and/or complement Wavefier to perform joint analysis (both online and offline) of data from different messengers/experiments
- Test case on MM data (both simulated and real)
  - LIGO-Virgo-KAGRA (GWs), CTAO (gamma rays), KM3NeT (neutrinos)
- Interoperability of science tools and data products
  - Including event data and instrument response functions



Credits: E. Cuoco, B. Patricelli, A. Iess, K. Graf, J. Schnabel, G. Hughes, A. Stamerra



## Software Metadata in OSSR

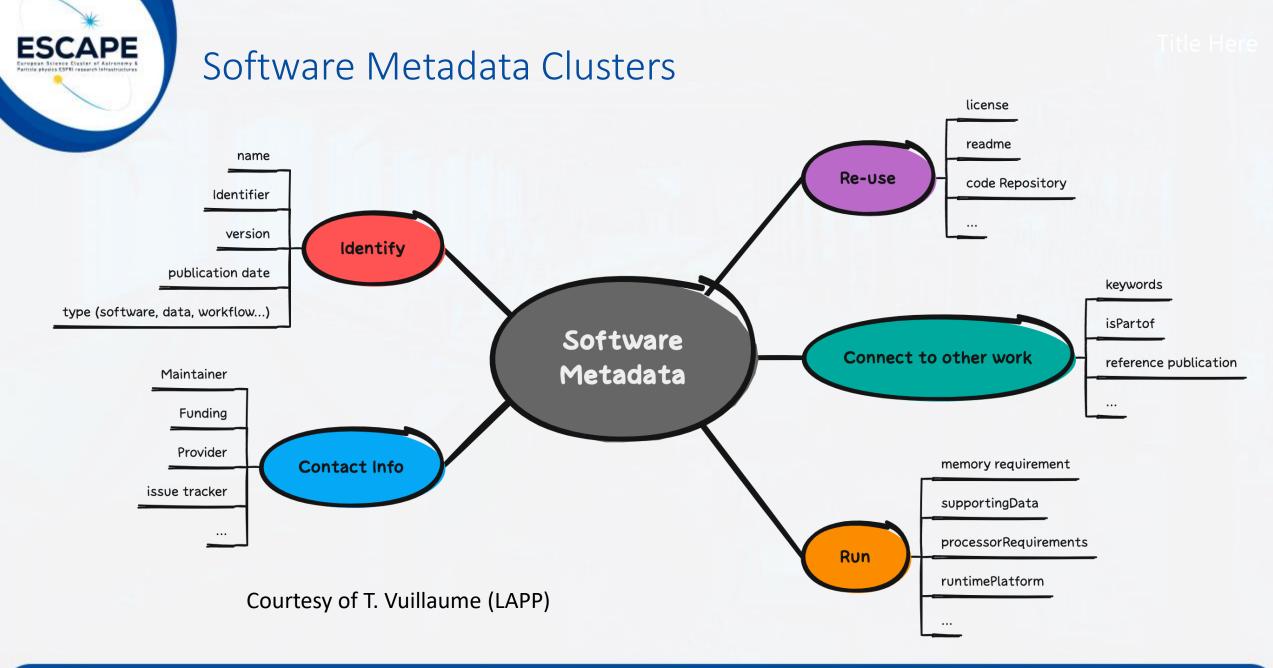
Our research software will follow FAIR principles (FAIRsFAIR)

FAIR principles establishes importance of metadata for software

- Software findable (FAIR principles)
- Workflows reproducible, provenance information
- Output in the second second
- Oiscoverable on software archives, public places, here: Zenodo
- OSSR extended metadata scheme of software:
  - Based on standard metadata schemas and software (e.g. codemeta)
  - Includes metadata on keywords and on the running environment (see next slide)
  - Not only data itself is the entry point, now also easier identification of workflows and corresponding software

Has become part of the ESCAPE recommendations and guidelines to CTAO and other ESFRIs

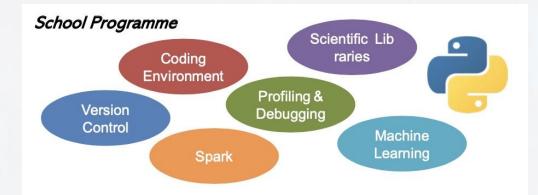


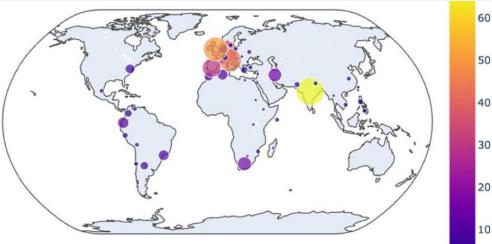




## OSSR is more than a Repository

- OSSR is also a forum and a school for both developer and user communities
  - Cross-ESFRI user community
  - Cross-ESFRI developer community
- Schools and Trainings as part of user support
  - Training of early career scientists is essential for the development of an open science system in the EOSC
  - Schools no limitations, completely free
  - All lecture contents openly accessible online
  - Driven by the ESFRIs and community-based software projects
- Cross-fertilization in developer forums and common developments
  - Improve interoperability, standardisation, and quality of open research software
  - Important that developers can discuss to each other in a thriving community
  - Common development examples: agnpy, GammaPy





Number of participants by country who answered the school survey.

From ESCAPE Summer School 2021, see references



ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

26/10/2022

### Conclusion

- CTAO contributed to the building of the OSSR (eOSSR library) and to the definition of software metadata
- CTAO participated in cross-ESFRI software developments and in the software training and schools
- OSSR shows great potential for increasing the ESFRI user base for the software and related analysis workflows
- OSSR provides a valuable resource for recommendations and guidelines and standardization
- OSSR is more than a repository and provides a forum for both developer and user communities to enhance the software quality and usage
- We look forward to further work together on these topics in future collaborations





ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 824064.

26/10/2022







OSSR work package for their active support and engagement

Kay Graf
Jutta Schnabel
Elena Cuoco
Cristiano Bozza
Mark Kettenis
Marjolein Verkouter
Thomas Vuillaume
Enrique Garcia

# Software Developments Agnpy: Cosimo Nigro, Javier Rico Gammapy: Axel Donath, Quentin Remy, Atreyee Sinha Concordia: Cristiano Bozza, Jose-Luis Contreras, Daniel Nieto Wavefier: Elena Cuoco, Alberto Iess, Barbara Porticelli, Antonio

Stamerra



## References: ESCAPE Summer Schools

### ESCAPE Summer School 2021

<u>https://escape2020.github.io/school2021/</u>

<u>https://indico.in2p3.fr/event/20306/</u>

ESCAPE Summer School 2022

<u>https://escape2020.github.io/school2022/</u>

<u>https://indico.in2p3.fr/event/26913/</u>

Output Content Cont

https://zenodo.org/record/5838436#.Y0b46i8RoUh



CAPE



# Thanks!





in company/projectescape

ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures **ESCAPE to the Future** 25-26 October 2022 Brussels, Belgium

# ESCAPE to the Future: OSSR – Open Software in the EOSC

## Kay GRAF

for the OSSR Team ESCAPE to the Future, Brussels, 25-26<sup>th</sup> of Oct. 2022



## OSSR to the Future

### Technical Developments:

- Extending metadata scheme
- Extending eOSSR library with advance search, additional development platforms and archives
- Support for integration with analysis platform and virtual research environment in EOSC-Future

### Sustaining OSSR:

- Repository infrastructure is sustainable by choice/design
- Interest group from members of onboarding group formed
- Continue the cross-fertilisation and co-operation in software development
- Maintenance is goal of the ESCAPE collaboration

### Enlarging the scope

- Engagement with HORIZON-INFRA-2023-EOSC-01-02 "Development of community-based approaches for ensuring and improving the quality of scientific software and code"
- Strive to become part of the EOSC Exchange layer
- Collection and preservation of all software and services generated in ESCAPE



# Thanks for your attention!





in company/projectescape